Module 2. Evaluation of Biometric Systems in Real Applications

Francesc Serratosa

Universitat Rovira i Virgili

Tarragona, Catalonia.

September 2018

francesc.serratosa@urv.cat

http://deim.urv.cat/~francesc.serratosa/

Introduction

The final objective of biometric systems is to increase the security of other systems. For example, if a system for identifying people through the iris is applied in airports the objective is to increase the security in that country. In this case, "the security in the country" is the system that wants to increase its security.

Therefore, it is essential to ensure that the biometric system itself is as secure as possible and have tools to evaluate this security. In this module we explain the tools that have been developed to determine the quality of biometric systems.

Also, in this module, we evaluate biometric systems and discuss several real biometric systems to show the importance of evaluating large systems. These biometric applications are of great social, economic and political importance with millions of people involved and can only be applied if they have very few biometric errors.

It is clear that it is acceptable to base the security of a system (for example a presidential election) on a biometric method if the biometric system has fewer errors than other non-automated or semi-automated methods.

Objectives

The objectives of this module are to explain the errors that may appear in a biometric system and explain the mechanisms that have been proposed to evaluate the goodness of biometric systems. We also show a few examples of real and large-scale biometric applications.

- Classify errors that may appear in a biometric system. Under which conditions do these errors appear.
- Evaluate a biometric application to determine its goodness. Discuss metrics for evaluating and comparing the goodness of biometric systems.
- Show a few examples of large biometric systems.

1 Errors of Biometric Systems

The promise on which biometric recognition is based is that given a new sample, the biometric system will always provide the correct decision, whether for verification (it is or it is not the person) or identification (the identification of the person is provided). In practice, a biometric system is a pattern recognition system that inevitably makes incorrect decisions. Therefore, it is essential to understand why a biometric system makes mistakes and model these errors to be able to determine their magnitude. And then discuss what kinds of errors we can find. The comprehensive treatment of biometric systems is detailed in (ISO/IEC 19795-2, 2007).

1.1 Reasons for errors in biometric systems

- **Limitation in the Information**: The distinctive, unvarying information contained in a biometric sample is inherently limited due to the intrinsic capacity of the identifier's signal and the biometric sensor. For example, there is less distinctive information in hand geometry than in fingerprints. Consequently, the samples of hand geometry can distinguish less identifications than fingerprints even in ideal conditions. The information may also be limited due to a poor presentation of the biometric feature to the sensor by the users or the acquisition of an inconsistent signal. The different ways of acquiring samples of a biometric feature limit the invariance of different samples from the same user.
- **Limitation in the Representation**: The ideal representation should be designed to restrict all invariance as well as the discriminatory information from the samples taken. Current modules for extracting characteristics, generally based on simplistic models of a biometric signal, fail to capture all the richness of the information in a real biometric signal. Thus, erroneous features are included and true features are excluded. Consequently, a part of the legitimate space of the samples cannot be represented by the biometric system, and thus representation errors appear.
- **Limitation in the Invariance**: Finally, given a representative scheme, the design of an ideal comparator should perfectly model the relationship of invariance through different samples of the same user (same identification), although the samples have been acquired under different conditions. Again, in practice (due to the inability to acquire a sufficiently large number of samples or the variance in the conditions when samples are captured) the comparator may not model the invariance relationships and thus errors appear in the comparator.

The challenge is to obtain a realistic and invariant representation of the biometric feature from new samples under uncontrolled (or almost uncontrolled) conditions, and then formally estimate the discriminatory information in the signal of the samples. This task is very difficult in a large-scale identification system in which the number of enrolled users can be enormous (more than 50 million).

1.2 Types of errors in biometric systems

This section details the errors that appear in the different stages of a biometric system (see Figure 10).

- **Errors in the Capture Module**: In completely automatic systems, the data are captured without supervision by any expert. These biometric systems generally use a *live-scan* device that automatically detects the presence of a biometric feature when it appears in the device's sensor. This type of capture can produce two types of errors: *Failure in Detection* (FD) and *Failure in Capture* (FC). Failure in detection appears when the biometric feature approaches the sensor but the sensor is not able to detect

its presence. Failure of capture occurs when the system realizes that there is a biometric feature but cannot capture the sample. Normally, the rate of these two failures is inversely proportional.

- **Errors in the Feature Extractor**: After capturing the sample, the system sends it to the feature extractor. If the image is of very low quality (careful, voice would be an acoustic signal), the extractor is not able to extract a coherent feature. This error is known as *Process Failure* (PF). Because the capture module and the feature extractor are used in the three basic processes (enrolment, verification and identification), they are usually joined into a single measure called *Acquisition Failure* (AF). A high percentage of acquisition failures in relation to the number of times that samples have been taken results in a significant decrease in the system's performance and frustration for users, which leads to rejection of the biometric system (reduction in acceptability). One way of reducing the percentage of acquisition failures is to allow the system to generate a set of features even if the image is bad, that is, the quality of these features is low. The problem is that then the comparison module has an additional load and comparisons can give erroneous outputs.
- **Errors in the Template Creation Module**: The template creation module can also make errors given a set of features of different samples. These failures appear when the features have been extracted in a very noisy situation and therefore there is little consistency between the samples. This failure is called *Failure in Enrolment* (FE) since the template is only generated in the enrolment process. Like acquisition failure, if enrolment failure is deactivated or very low quality limits are set, then there are many more errors in the comparison.
- **Errors in the Comparison Module**: The comparison module generates a result given a sample and a template. This result tends to have a value within the range of 0 to 1 and represents a probability or a distance. The possible failures of the comparison module depend on whether we are in a verification or identification process. Now we will describe them.

  In a **verification process**, after calculating the distance or probability, a threshold is applied, which can be modified externally, to arrive at a final decision. If the distance (or probability) is lower (or higher) than the threshold, then it is considered that the template and the sample come from the same individual. Otherwise, it is considered that they are from different individuals. In this process, there are four possible combinations, two of which generate errors:
    - The user is identified correctly and presents their biometric features to the system:
      - The system returns correctly that there is a match, that is, that the biometric features belong to the identification presented. There is no error and this is called *Correct Acceptance*.
      - The system wrongly returns that there is no match, that is, that the biometric features are not from the person with the identification presented to the system. This is a *False Non- Match or False Rejection*.
    - The user is falsely identified (for example, the user introduces the identification of another person who they know has special permissions) and presents their own biometric features to the system:
      - The system correctly returns that there is no match, that is, that the biometric features belong to another identification. There is no error and this is called *Correct Rejection* (CR).
      - The system wrongly returns that there is a match, that is, that the biometric features are from the person presenting the identification to the system. This is a *False Match or False Acceptance* (FM or FA).

In an **identification process**, there are six possible combinations, three of which generate errors and one is not possible.

- The person whose biometric features are being searched for has enrolled in the system (their template is in the database).
  - The system returns the identification of the person who is being searched for. There is no error: *Correct Acceptance,* (CA).
  - The system returns another identification. This is a *False Positive Identification* (FPI). What has happened is that there is a template of another person who by error has returned a shorter distance than the correct template.
  - The system returns that there are no templates with these biometric features. This is a *False Rejection* (FR). This case can only occur when the identification system has a threshold like in the verification system. This error can be eliminated by making the distance threshold less restrictive, that is, we increase its value.
- The person whose biometric features are being looking for has not enrolled in the system (their template is not in the database).
  - The system returns the identification of the person who is being searched for. This combination is not possible. If the identification has never been entered because the user has not enrolled, then the system can never return their identification.
  - The system returns another identification. This is a *False Negative Identification* (FNI).
  - The system returns that there are no templates with these biometric features. This is a *Correct Rejection* (CR). Like rejection error, this situation can only occur if there is an acceptance threshold. No template has returned a distance lower than the acceptance threshold. If the acceptance threshold is increased to stop rejection errors, then we might find that some correct rejections disappear.

1.3 Modelling the Errors

In the previous section, we detailed what kinds of errors can appear in a biometric system and the causes that generate them. In this section we make a more scientific study of these errors in a verification system and in an identification system.

In a Verification System

Let's suppose the template stored in the database of a person is **T** and the sample we want to verify is **I**. Also, let's suppose we have a similarity function (defined as the inverse of a distance) between a sample and a template S(**I**, **T**). S takes values within the range of 0 to 1. The larger S is, the more the sample resembles the template, that is, the higher the probability that they belong to the same person. Then we have two possible hypotheses:

H_0: $\mathbf{I} \neq \mathbf{T}$: The sample that we want to verify does not belong to the person who has generated the template.

H_1: $\mathbf{I} = \mathbf{T}$: The sample that we want to verify is from the person who generated the template.

The possible responses from the biometric system are:

D_0: There is no match. The system considers that they belong to different people.

D_1: There is a match. The system considers that they belong to the same person.

Considering the hypotheses and the system outputs, we find the errors that we mentioned in the previous section:

- *False Match,* FM: Also called type I error. The system returns D_1 when the hypothesis was H_0.
- *False Non-Match,* FNM: Also called type II error. The system returns D_0 when the hypothesis was H_1.

False Match Rate FMR is the probability of a type I error. Mathematically, FMR = Probability (D_1| H_0).

False Non-Match Rate FNMR is the probability of a type II error. Mathematically, FNMR = Probability (D_0| H_1).

To be able to evaluate the accuracy of a biometric verification system, it is necessary to collect a very high number of comparisons between samples and templates of the same person as well as a very large number of comparisons between samples and templates of different people. A biometric sample is represented with s. The set of first samples is called the *genuine distribution* and mathematically is represented by the distribution p(s|H_1). The set of second samples is called the *impostor distribution* and mathematically is represented by the distribution p(s|H_0). Therefore, we can define the reasons for the errors with the following functions:

$$FNMR = \int_0^t p(s|H_1)\,ds \text{ i } FMR = \int_t^1 p(s|H_0)\,ds$$

Where ***t*** is the acceptance threshold imposed by the system administrator.

Figure 1 shows the impostor and genuine distributions with respect to the matching score. In a real system, the samples that belong to the genuine distribution tend to have a greater similarity (or smaller distance, more to the right in the figure) than the samples that belong to the imposter distribution (with a greater distance, more to the left of the figure).

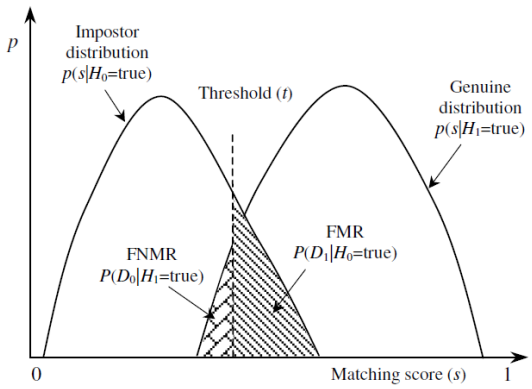


Figure 1. Distribution of impostor and genuine populations with respect to similarity (matching score).

There is a very large relation between the FMR (or False Acceptance Rate, FAR) and the FNMR (or False Rejection Rate, FRR) in each biometric system. In fact, as seen in the formulas, they both depend on the acceptance threshold ***t***. Therefore, in reality, we should write FMR(***t***) and FNMR(***t***). As shown in the figure, the FNMR(***t***) (or FRR(***t***)) is the area marked by the genuine distribution and the threshold ***t***. And the FMR(***t***) (or FAR(***t***)) is the area marked by the imposter distribution and also the threshold ***t***. If ***t*** is decreased to make the system more tolerant to the variations in inputs and noise, then the FMR(***t***) increases. However, if we increase ***t*** to make the system secure then we increase the FNMR(***t***). The system administrator cannot know in advance

where the system will be deployed or the users' response to it. Therefore, it is initially difficult to set the threshold ***t***. The following two functions have been defined to determine the goodness of a verification system regardless of the threshold:

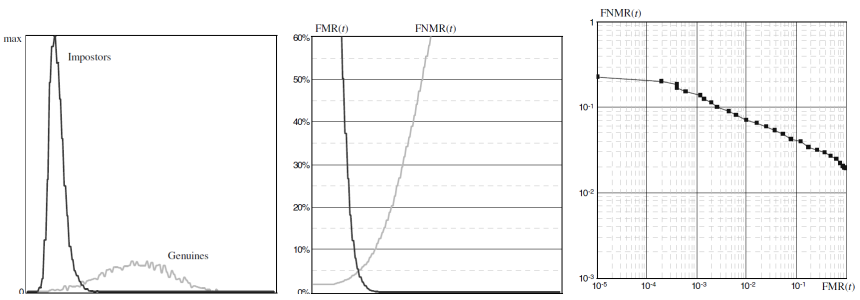
- Receiver Operating Characteristic (ROC): The ROC is a curve in a two-dimensional plane marked by the points FMR(***t***) and 1- FNMR(***t***) for various values of ***t***. The value 1- FNMR(***t***) is called the *power of the test* or *the goodness of the test*. This graph shows the FMR with respect to the power of the test and the goodness of the test
- Detection-Error Trade-off (DEC): The DEC is a curve like the ROC but marked by the points FMR(***t***) and FNMR(***t***). The DEC is interesting for showing the relationship between the two types of errors, since the objective is to minimize the two errors.

In addition to these graphs, there are three global indices for analysing a biometric system. When we receive information from a biometric system from a person in the company, it is important to use or consider these indices carefully because they have usually been carried out scientifically, but with a database controlled by the system's own developers. The indices are:

- *Equal-Error Rate* EER: Indicates the error rate for all threshold values that the FMR is equal to the FNMR. EER = FMR(***t***) such that FMR(***t***) = FNMR(***t***) for all ***t***.
- *Zero FNMR*: This is defined as the FMR in which there are no False Matches.
- *Zero FMR*: This is defined as the FNMR in which there are no False Non-Matches.
- *Separability*: If we assume that the genuine and imposter populations generate normal distributions (Gaussian distributions), then we can analyse how separate they are, or in other words, the little overlap that we find between the two populations. The more overlap, the more errors will be generated in the recognition process.

$$S = \frac{\left\| \hat{x}_{impostor} - \hat{x}_{genuine} \right\|}{\sqrt{\frac{\sigma^2{}_{impostor} + \sigma^2{}_{genuine}}{2}}}$$

Figure 2 shows the results of an algorithm for comparing the fingerprints presented in the Fingerprint Verification Competition (FVC) in 2002. The FVC is a competition in which companies or research centres can send their algorithms that are already compiled (the source code is not sent) to evaluate its functionality and goodness. The data shown were calculated with 2800 pairs of genuine fingerprints (they belonged to the same finger) and 4950 pairs of imposters (belonging to different individuals).



(a) (b) (c)

Figure 2. a) Genuine and imposter distributions with respect to similarity (match score). b) Evolution of the error percentage of the FMR and FNMR errors relative to the threshold. c) The DEC curve obtained with the same data as computed in Figure b).

Figure 3 shows the percentage of FMR and FNMR errors with respect to the threshold. The point where the EER is defined as well as the Zero FNMR and Zero FMR values are also shown. Notice that the Zero FNMR value (or Zero FMR) is located exactly at the point where the graph of the FNMR (or FMR) has a value of 0.

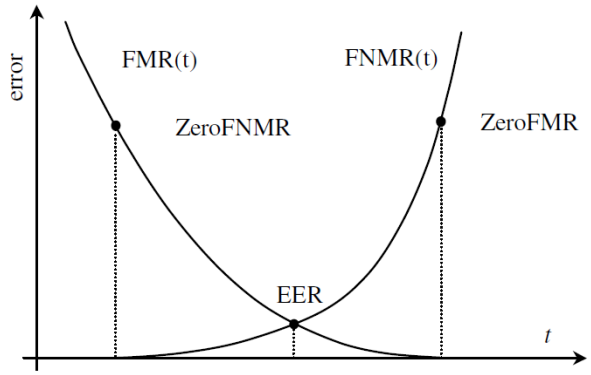


Figure 3. Example of obtaining the overall Zero FNMR, Zero FMR and EER values with the curves of the error rates relative to the similarity threshold.

The accuracy requirements of a biometric verification system depend greatly on the application. For example, in forensic applications used to identify criminals, what we want is to make sure we don't not identify a criminal even if there is the risk of having to manually examine a lot of potentially false matches identified by the system. This implies that what we want is a high FNMR and therefore, we will put a low similarity threshold. Another extreme would be a high security access control. The main objective is that impostors do not enter. In this case, we want the FMR to be high. Clearly, if we set the similarity threshold very high, then we will lower the FMR, but this will imply that sometimes there will be authorized people who will not be given access. Figure 3 clearly shows this concept.

In an Identification System

Suppose a sample you want to identify is compared to N templates in the database and let's also assume that these comparisons are independent of each other. In identification systems, as we have already seen, three types of errors are defined that are related to false matches and false non-matches:

- *False Positive Identification, FPI*: This is directly related to the False Non-Match. We have the correct sample and the template exists but the system is not able to find the correct labelling.
- *False Rejection, FR*: This is like the false positive identification, therefore, it is also related directly to the False Non-Match, since we have the correct sample and the template exists but the comparator returns a similarity below the threshold (or a distance greater than the threshold).
- *False Negative Identification, FNI*: this is related to the False Match. In this case we do not have the correct template and the system returns an incorrect match and that the similarity is greater than the similarity threshold.

The reason for the *False Positive Identification-Error Rate* FPIR includes the reason for the first two errors mentioned: False Positive Identification and False Rejection. This error depends on the number of templates ***N*** and is defined $FPIR_N = 1 - (1 - FMR)^N$. These errors appear when the sample is erroneously matched with one or more templates in the database. Therefore, $FPIR_N$ is calculated as 1 minus the probability that no false match will be made with any of the templates. The expression (1 - FMR) is the probability that the sample will not be falsely matched with one of the templates in the database. If the FMR is very small, then this general expression can be approximated by $FPIR_N \approx \boldsymbol{N} \cdot FMR$. And in this way we can establish that the probability of false positive identification increases linearly with the size of the database. This approach is based on only considering the first term of Newton's binomial. If we wanted a more accurate approach, we could use the first two terms of Newton's binomial and the expression would be: $FPIR_N \approx \boldsymbol{N} \cdot FMR - \frac{N \cdot (N-1)}{2} \cdot FMR^2$. With this second approach, the value that is obtained is a little smaller since it has one more remaining term.

The *False Negative Identification-Error Rate* FNIR is easier to calculate since it is considered to be exactly the same as the FNMR; FNIR = FNMR. This is because the probability of a false negative identification when we look for the template in the ***N*** templates of the database is the same as the FNMR in the verification mode.

2 Evaluation of a Biometric System

The goodness of a biometric system depends dramatically on a lot of variables: The composition of the population (employment, age, sex, demography, race, etc.), the environment, the way of doing the tests as well as other specific restrictions of the application. In an ideal situation, we would like to characterize the performance in a model that is independent of the application. Therefore, it is possible to predict the performance of a real application. Rigorous modelling techniques have been applied to characterize the data acquisition and the comparison process. With these techniques, it has been possible to extrapolate the results obtained in the laboratory as though they were real applications, obtaining quite good results. Nowadays, comparative evaluations are being carried out with small databases. The clearest examples are the Fingerprint Verification Competition FVC (which we have just mentioned) and Iris Verification Competition IVC. The results obtained in these competitions can determine whether a system can be marketed or not. Due to the importance of being able to evaluate the accuracy of biometric systems, three types of evaluations can be defined:

- **Evaluation of the Technology**: The objective is to evaluate the quality of the algorithms given a specific technology. The whole system is not evaluated but rather algorithm by algorithm. All algorithms are compared given the same sensors, database and any aspect that could affect the results. The database is divided into two parts. Normally, all data are generated at the same time and the partition is made randomly. The first part consists of the *Learning Database*. It forms the part of the data that users can use to be able to tune the algorithm and obtain the maximum performance. The second part comprises the *Test Database*. This is the part of the data that the evaluators use to make the final tests. The participants are not able to use or view it before doing the tests. Because the data are made available to the entire scientific community, the experiments can be repeated later. Some books on biometrics include DVDs with these data. It is a *repeatable assessment*.
- **Evaluation of the Scenario**: The objective of this type of assessment is to determine the complete performance of the entire system in a laboratory prototype or in an application simulator. The test is carried out in a complete system but under controlled conditions, although it attempts to simulate a situation in the real world. The comparison is always carried out with the same biometric sensors and the same population. It is a *repeatable assessment*.
- **Evaluation of the Performance**: The objective of this evaluation is to determine the performance of the complete system in a real situation of a specific environment and a specific population. It is a *non-repeatable assessment* because there may be undocumented or unknown parameters. There is no initial database.

3 First Large Real Applications

The following list gives a few examples of applications developed on a large scale. It does not attempt to be an exhaustive list but rather provides a few examples to show that biometrics are being applied and have been applied for some time now in real problems around the world. In addition, we have selected applications that are not related to access control or security, which are the most common.

South Africa *Fingerprint verification*

The first large-scale biometric application using fingerprints was the distribution of pensions in rural areas of South Africa. In 1990 each pensioner registered their fingerprints. They were stored in a personal card and verified before their pension was given to them to ensure that the person who was carrying the card was actually its owner. This system considerably reduced fraud.

Mexico *Fingerprint verification*

In Mexico, the Federal Electoral Institute installed 2000 biometric control devices to verify voter identity cards. The purpose was not to identify the voter by name, but to verify that he or she, despite their name, had the right to vote. Apparently, the operation was a success.

Uganda *Face verification*

To fight against electoral fraud, the Ugandan president decided to have a face recognition system installed at the polling stations for the general elections in June 2001. In two months, 11 million voters were photographed to create a database only for the elections. This system is still being used.

Malaysia *Fingerprint verification*

Since 2001, a biometric identification card has been issued to every inhabitant of Malaysia over 12 years old, the *MyKad*, which contains information like date and place of birth, sex, parents' names, ethnicity of origin, religion, a photograph and fingerprints. This card has many purposes as it serves as a driving license, passport, electronic payment card and also contains emergency medical information.

http://en.wikipedia.org/wiki/MyKad

Afghanistan *Iris verification*

The High Commissioner for Refugees (HCR) used biometrics in 2003 to help Afghan families return to their country after a long stay in Pakistan. HCR staff photographed the iris of the people that could possibly return. When they were ready to return, they were identified before they were given the money for transport, food coupons and basic needs. The agency stated that in this way they saved millions of dollars by preventing identity fraud.

Australia *Iris verification*

In 2003, a biometric system was tested in Australia to dispense methadone. The iris of drug addicts who chose to participate in the programme was photographed for their subsequent identification at the pharmacies participating in the study, where patients would receive the exact dose prescribed by the doctor. The use of this control system is particularly useful in large pharmacies where pharmacists do not know all their clients.

Europe *Fingerprint verification*

In an attempt to reduce the many requests for political asylum in the different countries of Europe, the European Community created a centralized database, *Eurodac*, which contains the fingerprints of all asylum seekers. Before January 2003, when the system became operational, it was estimated that 80% of 500,000 annual applications were requested in several countries, which has now been reduced to 11% of 280,000 applications. The *Eurodac* database cannot be matched to other databases.

http://en.wikipedia.org/wiki/Eurodac

India *Fingerprint verification*

Certain religious ceremonies in the temple of Tirumala in India can attract 150,000 pilgrims a day. Authorities have adopted a biometric system to help manage the crowd. Pilgrims register their fingerprints in advance. On the day of the ceremonies, they are identified by their fingerprints and allowed to enter the temple.

Japan *Fingerprint and face verification*

In Kyoto an experiment was carried out to allow the elderly to benefit from social services without leaving their homes. Registered people connect with city council workers and identify themselves by showing their face to a webcam and putting their finger on a fingerprint sensor.

Thailand Indonesia *Fingerprint and DNA verification*

To identify the bodies after the tsunami in December 2004, the experts collected DNA samples that were then compared with the DNA of families looking for a missing relative, as well as the fingerprints of missing people who had their fingerprints registered in identification documents.

Texas, USA *Fingerprint verification*

Medicaid is an American programme that provides health benefits to people with low incomes. To provide better protection of patient medical information and reduce fraud, the state of Texas has been testing a pilot project since 2004. The fingerprints of people eligible for *Medicaid* are stored in their Medicaid card and are checked before they receive benefits.

http://en.wikipedia.org/wiki/Medicaid

Annex

Obtaining false acceptances, false rejections and DEC from the similarity matrix

This section describes how to obtain a DEC from the similarity matrix. That is, the matrix in which the rows and columns represent specific enrolments of biometric features and the cells represent the similarity between the biometric features. Table 1 shows an example of a similarity matrix. Suppose we have a database with three people and each person has made three recordings. It is normal for a user to make three recordings when they enrol to ensure there is more information about their biometric features. These 9 recordings are represented by the columns. On the other hand, suppose five people have shown their biometric features to try to access the database. The first three people are those who enrolled but the last two have not enrolled. These five verification attempts are represented in the five rows.

	Person 1			Person 2			Person 3		
Person 1	**3**	**3**	**4**	3	2	3	3	0	1
Person 2	3	2	4	**1**	**3**	**4**	2	3	6
Person 3	1	4	1	2	1	2	**3**	**8**	**5**
Person 4	3	1	9	2	6	3	5	0	7
Person 5	2	1	3	3	0	0	1	2	2

Table 1. Similarity matrix. Three people registered 3 times. 5 attempts of verification.

Genuine authentication attempts are the values marked in bold type. Person 1, person 2 and person 3 say that they are who they really are and the comparison with their biometric features is carried out. Note that it has to be considered that there are only three genuine attempts and not nine despite the fact that nine comparisons have been made. The rest of the values are fraudulent identification attempts. The person presents their biometric features to the system but says that they are another person. There are 3x5 identifications minus 3 genuine identifications = 12.

The result of a verification is binary: "It is accepted that the person is who they say they are" or "It is not accepted that the person is who they say they are". This decision is made by applying a threshold (set by the system administrator) and checking the three registrations of the person that the user says they are. Only when one of the three comparisons is above the threshold, do we then consider that we have a correct request: "It is accepted that the person is who they say they are". On the other hand, it is necessary that the three comparisons are below the threshold for us to believe that the petition is incorrect: "It is not accepted that the person is who they say they are".

Suppose the threshold is 4.5. Table 2 shows the results of the verification.

The errors are marked in bold. The false acceptances are the cells with bold and underlined values. That is, impostors for whom one of the three results of the comparison was above the threshold of similarity. The false rejections are the cells with bold and crossed-out values. That is, genuine verifications that the threshold is above the three values.

The False Match Rate or FMR is calculated as the number of false acceptances (bold and underlined cells) divided by the imposter population. With a threshold of 4.5 the value is: FMR = 4/12 = 1/3.

The False Non-Match Rate or FNMR is calculated as the number of false rejections (cells in bold and crossed out) divided by the genuine population. With the threshold at 4.5 the value is: FNMR = 2/3.

Threshold 4.5	Person 1	Person 2	Person 3
Person 1	~~**Different person**~~	Different person	Different person
Person 2	Different person	~~**Different person**~~	**The same person**
Person 3	Different person	Different person	The same person
Person 4	**The same person**	**The same person**	**The same person**
Person 5	Different person	Different person	Different person

Table 2. Result of the verification given the similarity matrix of Table 1 and threshold 4.5.

To draw the DEC, we calculate the FMR and the FNMR values for various threshold values, for example in this case: 0.5; 1.5; 2.5; 3.5; 4.5; 5.5; 6.5; 7.5; 8.5; 9.5. With the 10 pairs of FMR and FNMR values obtained, we can draw the DEC.

Summary

In this module we have described how to evaluate biometric systems. First, we looked at the errors that may appear in biometric systems. Next we saw that there are global metrics and that there are graphical metrics like the ROC and DEC. It is essential to evaluate biometric systems because biometrics is a highly-applied science.

We also described examples of applications on a large scale. These examples should be seen simply as a sample of the possibilities that biometrics provides for person recognition. Other systems have been put into operation and new ones appear.

Activities

1. **Errors in biometric systems**
   The reasons for errors in biometric systems are classified into three different types of "limitations". Describe and summarize them.
2. **Errors in the modules of biometric systems**
   Given the above limitations, there are four types of errors in the modules (or processes) of a biometric system. Describe the modules as well as the possible errors.
3. **General errors in biometric systems: Verification**
   Biometric systems have a threshold (it is usually not visible to the user or even the system administrator) based on which it is considered that given a comparison, the two samples come from the same individual or not. Relate this threshold to the general errors of biometric verification systems: False Match and False Non-Match.
4. **General errors in biometric systems: Identification**
   Do the same exercise as above but for the possible errors of biometric identification systems: False Positive Identification, False Rejection and False Negative identification.
5. **Modelling errors**
   Explain what is a genuine distribution and what is an impostor distribution and how they are modelled mathematically.
6. **Verification. Modelling errors**
   Given the genuine values of similarity: {3, 3, 5, 5, 6, 6, 6, 7, 9} and the impostor values of similarity: {1, 2, 2, 3, 3, 3, 4, 4, 5}, draw the distribution function of FNMR and FMR (Figure 1). What is the threshold value that makes the general error (FNMR + FMR) lowest? What is the lowest threshold value for FNMR = 0? With this threshold, what is the value of FMR? What is the highest threshold value for FMR = 0? At this threshold, what is the value of FNMR?
7. **Verification. Graphics representing goodness**
   Draw ROC and DEC distributions of the population of the previous exercise. Suppose the following 10 thresholds: {0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5}
8. **Verification. Population study**
   Given the populations of the two previous exercises, what are the EER, Zero FNMR, Zero FMR and separability of these populations?
9. **Evaluation of biometric systems**
   Describe the three types of evaluations: Technology, Scenario and Performance.
10. **DEC and ROC Graphs**
    Draw the DEC and ROC given Table 1 in the Annex section.
11. **Real applications**
    Search for real applications in which biometric techniques are applied and group them according to specific biometric features.

Bibliography and References

- Anil K. Jain, Patrick Flynn and Arun A. Ros (editors), "Handbook of Biomtrics", Editorial Springer, Year 2008, ISBN 978-0-387-71040-2
- David D. Zhang, "Automated Biometrics. Technologies and Systems", Editorial: Kluweer Academic Publishers, Any 2000, ISBN 0-7923-7856-3
- Anil Jain, Ruud Bolle i Sharath Pankanti, "Biometrics. Personal Identification in Networked Society", Editorial Kluweer Academic Publishers, Any 1999, ISBN 0-7923-8345-1
- James Wayman, Anil Jain, Davide Maltoni i Dario Maio (editors), "Biometric Systems. Technology, Design and Performance Evaluation", Editorial Springer, Any 2005, ISBN 1-85233-596-3
- Samir Nanavati, Michael Thieme i Raj Nanavati, "Biometrics. Identity Verification in a Networked World". Editorial Wiley Computer Publishing, Any 2002, ISBN 0471-09945-7
- Arun A. Ross, Karthik Nandakumar i Anil K. Jain, "Handbook of Multibiometrics", Springer, Any 2006, ISBN 0-387-22296-0

Acronyms

CA: Correct acceptance

CR: Correct rejection

DEC: Detection-Error Trade-off.

EER: Equal-Error Rate.

FM: False Match

FNM: False Non-Match

FA: False Acceptance

FR: False Rejection

AF: Acquisition Failure

FC: Failure in Capture

FD: Failure in Detection

FE: Failure in Enrolment

FMR: False Match Rate

FNMR: False Non-Match Rate

PF: Process Failure

FNIR: False Negative Identification-Error Rate.

FPMR: False Positive Identification-Error Rate.

FPI: False Positive Identification.

FNI: False Negative Identification

ROC: Receiver Operating Characteristic.